Insect Biochemistry and Molecular Biology 48 (2014) 100–109

Contents lists available at ScienceDirect

Insect Biochemistry and Molecular Biology

journal homepage: www.elsevier.com/locate/ibmb

Effects of different dietary conditions on the expression of trypsin- and chymotrypsin-like protease genes in the digestive system of the migratory locust, *Locusta migratoria*

Jornt Spit*, Sven Zels, Senne Dillen, Michiel Holtof, Niels Wynant, Jozef Vanden Broeck

Department of Animal Physiology and Neurobiology, Zoological Institute K.U.Leuven, Naamsestraat 59, B-3000 Leuven, Belgium

ARTICLE INFO

Article history:
Received 25 January 2014
Received in revised form
6 March 2014
Accepted 6 March 2014

Keywords:
Diet
Digestion
Inhibitor
Insect
Proteolytic
Serine
Protease
Starvation

ABSTRACT

While technological advancements have recently led to a steep increase in genomic and transcriptomic data, and large numbers of protease sequences are being discovered in diverse insect species, little information is available about the expression of digestive enzymes in Orthoptera. Here we describe the identification of *Locusta migratoria* serine protease transcripts (cDNAs) involved in digestion, which might serve as possible targets for pest control management. A total of 5 putative trypsin and 15 putative chymotrypsin gene sequences were characterized. Phylogenetic analysis revealed that these are distributed among 3 evolutionary conserved clusters. In addition, we have determined the relative gene expression levels of representative members in the gut under different feeding conditions. This study demonstrated that the transcript levels for all measured serine proteases were strongly reduced after starvation. On the other hand, larvae of *L. migratoria* displayed compensatory effects to the presence of Soybean Bowman Birk (SBBI) and Soybean Trypsin (SBTI) inhibitors in their diet by differential upregulation of multiple proteases. A rapid initial upregulation was observed for all tested serine protease transcripts, while only for members belonging to class I, the transcript levels remained elevated after prolonged exposure. In full agreement with these results, we also observed an increase in proteolytic activity in midgut secretions of locusts that were accustomed to the presence of protease inhibitors in their diet, while no change in sensitivity to these inhibitors was observed. Taken together, this paper is the first comprehensive study on dietary dependent transcript levels of proteolytic enzymes in Orthoptera. Our data suggest that compensatory response mechanisms to protease inhibitor ingestion may have appeared early in insect evolution.

1. Introduction

Serine proteases such as trypsins and chymotrypsins that act in the midgut are the main proteolytic digestive enzymes in insects belonging to the orders of Lepidoptera and Orthoptera. They catalyze the breakdown of proteins to generate free amino acids necessary for insect growth and development. Therefore, proteases are considered potential candidates for the development of pest control strategies based on anti-nutritional protease inhibitors (Gatehouse, 2011).

Protease inhibitors are commonly recognized as a substantial part of the natural defense mechanisms in plants upon herbivory. They are present in many seeds and/or can be induced after insect feeding. However, the intimate association between insects and their host plants has led to distinct physiological responses of insects to the ingestion of dietary inhibitors (Mello and Silva-Filho, 2002). Possible mechanisms of insect resistance to protease inhibitors include overproduction of existing proteases to outnumber inhibitors present in the digestive tract (Brioschi et al., 2007) or the induction of specific inhibitor insensitive enzymes (De Oliveira et al., 2013; Jongsma et al. 1995). It has been suggested that these flexible, compensatory response mechanisms are often possible because of the large repertoires of proteolytic enzyme encoding genes that are present in many insect species. The occurrence of multiple isoforms may provide adaptive advantages for insects feeding on plants containing inhibitors. Therefore, it is not surprising that during the past decades, a large number of protease encoding genes were identified to be present in the digestive system of many different insects (Bown et al., 1997; Gatehouse et al., 1997; Ge et al., 2012; Marshall et al., 2008; Oliveira-Neto et al., 2004; Pedra et al., 2003; Prabhakar et al., 2007; Simpson et al., 2007).

* Corresponding author.
E-mail addresses: jorntspit@hotmail.com, jornt.spit@bio.kuleuven.be (J. Spit).

http://dx.doi.org/10.1016/j.ibmb.2014.03.002
0965-1748/

At present, little is known about protease inhibitor (PI) induced responses in Orthoptera. In addition, the exact mechanisms of these physiological responses, as well as the fundamental pathways that are responsible for the regulation of digestion-related processes in general, remain unclear. Recently, we were able to show that similar responses as observed in coleopteran and lepidopteran pest species are also present in the desert locust *Schistocerca gregaria* (Spit et al., 2012). In the present paper, we report on PI induced response mechanisms in the migratory locust, *Locusta migratoria*. Similar to the desert locust *S. gregaria*, the migratory locust *L. migratoria* can switch from a solitary phase to a swarming gregarious phase due to changing environmental conditions (Kang et al., 2004; Verlinden et al., 2009). At irregular time intervals, these swarms form a serious agricultural threat to large parts of Africa, Asia and Europe. Thorough knowledge of their digestive physiology has potential to contribute to the development of novel insect pest control strategies.

Because of their relatively large size, locusts have been an important research model in physiological studies for the past decades. However, compared to several other pest insects (e.g. Lepidoptera, Coleoptera), only limited sequence information of digestive enzymes was available for Orthoptera. Until very recently, only two mRNA sequences encoding a serine protease were described in locusts; a partial mRNA sequence encoding a trypsin-like protease in *L. migratoria manilensis* (Wei et al., 2007) and a transcript coding for a serine protease related protein in *S. gregaria* (Chiou et al., 1998). However, both sequences are almost certainly not directly involved in protein digestion in the midgut, since they were shown to be involved in molting and vitellogenesis, respectively. Nevertheless, recent sequencing efforts have led to an tremendous increase in sequence data available for locusts (Badisco et al., 2011; Ma et al., 2006; Zhang et al., 2012).

In this context, we aimed at identifying the different serine proteases present in the existing EST databases of *L. migratoria* and assessing their possible role in the PI induced compensation mechanisms that take place in the midgut. A total of 20 putative serine proteases were identified and by phylogenetic analysis of these sequences we showed the existence of three distinct evolutionary clusters. Transcript levels for different representative members of these clusters were examined by quantitative real time RT-PCR in different feeding conditions: starved and after soybean bowman birk (SBBI) and soybean trypsin inhibitor (SBTI) ingestion.

2. Materials and methods

2.1. Rearing of animals and sample collection

Locusts (*L. migratoria*) were reared under crowded conditions with controlled temperature (32 ± 1 °C), light (14 h photoperiod) and relative humidity (40–60%). Locusts were fed daily with grass. Fifth instar larvae were developmentally synchronized at the day of the 5th larval molt (Day 0). In order to obtain biologically active digestive peptidases, midguts and gastric caeca of 5th instar larvae were dissected, cleaned and incubated during 30 min in Ringer solution. Subsequently, the tissues were removed and the solution containing secreted enzymes was used for further handling. For extraction of total RNA, tissues were dissected in Ringer's solution and immediately transferred to liquid nitrogen. Samples were stored at −80 °C until further processing.

2.2. Feeding experiments

Locusts used in feeding experiments were placed in separate cages and fed on an agar based artificial diet containing 2.4% Wesson salt mixture, 0.5% linoleïc acid, 0.6% cholesterol, 18.8% Vanderzant vitamin mixture, 50% cellulose, 14% dextrin, 8.1% caseïne, 2.8% peptone and 2.8% albumin. Experimental groups were formed by random selection ($n > 25$/condition). Animals were individually weighted during the course of the experiments. After synchronization all groups were reared on control diet, receiving freshly prepared artificial diet daily. From day 4 after synchronization one experimental group was transferred to a diet containing plant derived PI (SBTI, SBBI: 1% of total protein content). A second group was starved for the same period of time. Both groups were compared to a control condition that was kept on control diet. At different time points (30 min, 2 h, 4 h, 48 h) after receiving the artificial diet, midgut and caeca samples were taken for RNA extraction and activity assays. Tissues for RNA extraction were dissected in at least 3 pools of 5 individuals for every condition.

2.3. Proteolytic activity and inhibitor sensitivity

In vitro assays were performed to measure proteolytic activity and/or inhibitory potency of different PI. Azocasein was used, providing a measure to assess all protease activity. Solutions containing secreted locust midgut enzymes were pre-incubated with buffer (control) or PI (total volume 100 μl) for 10 min at 32 °C. 100 μl azocasein (1%) was added and incubated at 32 °C for 30 min. The reaction was terminated by addition of 75 μl of 10% trichloroacetic acid. After centrifugation (16,000*g*, 10 min, 4 °C), 14 μl of 5 M NaOH was added to 90 μl of the supernatant and absorbance was measured at 405 nm in 3 technical replicates. Percentage inhibitory activity was calculated as Inhibitory activity (%) = 100 − [AbsPI/min AbsControl/min] × 100.

Inhibitory potency of SBBI and SBTI was tested over a range of concentrations. Final inhibitor concentration for AEBSF and cystatin was respectively 10 mM and 50 μM. In addition, the selective substrates N-benzoyl-phe-val-arg-p-nitroanilide (BPVApNA) and N-succinyl-ala-ala-phe-p-nitroanilide (SAAFpNA) were used to measure the trypsin and chymotrypsin activity, respectively. Forty microliters of enzyme solution were combined with 10 μl of buffer and pre-incubated at 32 °C for 10 min. Next this mixture was added to 50 μl of 1 mM substrate and absorbance was measured at 405 nm over a period of several minutes. The slope of the linear curve is a measure for the peptidase activity in the sample. All inhibitors and substrates were purchased from Sigma.

2.4. In silico detection and phylogenetics

A tBLAST-N search was used to search the EST database from *L. migratoria*, using multiple insect trypsin and chymotrypsin sequences as a query (Ma et al., 2006). Resulting EST sequences were assembled into contigs using the DNA Baser software. Sequences obtained were translated into the corresponding amino acid sequence (Prosite, ExPASy). Translated proteins were aligned with MAFFT alignment software (Katoh et al., 2005), using the BLOSUM matrix and manually verified. Signal peptide predictions were made using SignalP 4.1 (Petersen et al., 2011). Phylogenetic analysis was performed using MEGA version 5.2 (Tamura et al., 2011) with aligned amino acid sequences that were manually trimmed to obtain regions with the highest homology (trimmed alignment for Locust sequences in Supplementary Fig. S1). Maximum likelihood trees were constructed using a WAG substitution model and tested by the bootstrap method, using 500 replications. All gaps were treated as missing data.

2.5. RNA extraction and cDNA synthesis

The Lipid tissue extraction kit (Qiagen) was utilized to extract RNA from dissected tissues. DNase treatment was performed to

remove traces of genomic DNA contamination. Quality and concentration of the extracted RNA were assessed using a Nanodrop spectrophotometer. Equal quantities of RNA were used as template to produce cDNA. cDNA synthesis was performed using the Superscript III reverse transcriptase kit (Invitrogen), and random hexamer primers (Invitrogen) and dNTPs (Roche), following the manufacturer's protocol.

2.6. Quantitative real time RT-PCR

Primer express software (Applied Biosystems) was used to design quantitative real time (RT-)PCR primers. All primer sequences are displayed as Supplementary Data in Table S2. The primers were validated with a standard curve based on a serial dilution of cDNA to determine the primer annealing efficiency. A dissociation protocol was performed to detect the presence of primer dimers and production of a single PCR product. For all transcripts, only a single melting peak was found. Each qPCR reaction was performed in duplicate and contained 10 μl SYBR green solution (Invitrogen), 1 μl of 10 μM of both forward and reverse primer (Sigma), 3 μl milliQ water and 5 μl cDNA. The PCR reaction was performed in a 96 well plate and analyzed by the StepOne System (ABI Prism, Applied Biosystems). Relative expression levels were calculated using the delta delta Ct method (Livak and Schmittgen, 2001). To correct for sample to sample variation, expression was normalized against *rp49* and *rps13*, the two most stably expressed reference genes, as determined with geNorm (Vandesompele et al., 2002).

2.7. Statistical analysis

Student's *t*-tests and Mann–Whitney *U* tests were used to determine significant differences in weight gain and protease activity and sensitivity respectively. Significant differences in gene expression were determined by ANOVA. For the feeding experiment, means were calculated from two independent experiments. All statistical analyses were performed using GraphPad Prism 5.

3. Results

3.1. Digestive proteolytic activity in 5th instar larvae is elevated after protease inhibitor ingestion while inhibitor sensitivity remains unaffected

The activity of secreted digestive enzymes of 5th instar larvae was tested in the presence or absence of different inhibitors. Our experimental data demonstrated that the chemical serine protease inhibitor AEBSF is capable of inhibiting approximately 90% of the enzyme activity. Cystatin, an inhibitor of cysteine proteases, inhibited the remaining proteolytic activity with an average inhibitory potency of 11% (Fig. 1). This confirms the predominant use of serine proteases as digestive enzymes in *L. migratoria*, with only a minor contribution of cysteine proteases to the total proteolytic activity. To further specify, the solution containing secreted gut enzymes was capable of hydrolyzing both BVPApNA and SAAFpNA, specific substrates for trypsin and chymotrypsin enzymes respectively (Supplementary Fig. S4). Therefore, a combination of two plant derived protease inhibitors, SBBI and SBTI, both known to be active against both insect trypsins and chymotrypsins, was tested over a range of concentrations to determine the strength of inhibition. A combination of these inhibitors was capable of inhibiting nearly all detectable serine protease activity (Fig. 1). Since only 6 μM was needed for maximal inhibition, this mixture constituted a valuable candidate for *in vivo* feeding trials.

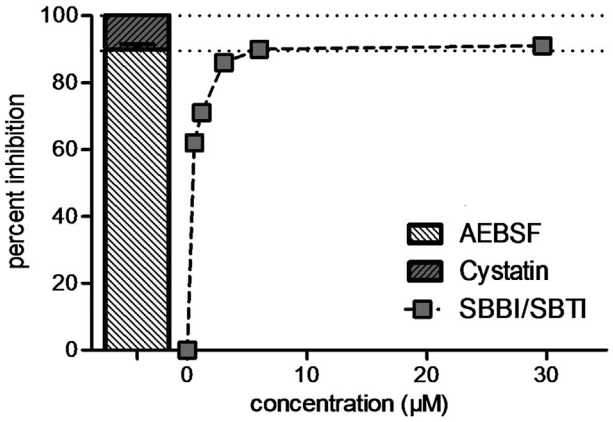

Fig. 1. Inhibitory strength of a combination of SBBI and SBTI compared to the chemical serine protease inhibitor AEBSF and the cysteine protease inhibitor cystatin. Inhibitory potency is calculated as percentage of a control measurement, without adding inhibitor, determined as 100% activity. Cumulative inhibitory potency of AEBSF (10 mM) and cystatin (50 μM) is presented. SBBI/SBTI potency against secreted midgut enzymes (filled squares) was tested over a range of concentrations. Data are presented by mean values ($n = 3$).

Next, the proteolytic activity in the midgut was quantified for locusts that had received a diet containing SBBI and SBTI (1% of total protein content) for three consecutive days and compared with control locusts that had received the same diet without any PI. During the course of the experiment, individual larvae were weighted, yet no clear effect on weight gain or development was observed, suggesting that *L. migratoria* larvae could readily adapt to the PI mixture. In addition, proteolytic activity was determined for larvae that were starved for three days. As shown in Fig. 2A, the proteolytic activity was more than doubled after feeding on the artificial diet containing inhibitors, while it was approximately half in starved animals. When the activity was assessed in the presence of SBBI and SBTI, only a negligible change in sensitivity was observed after PI ingestion (Fig. 2B).

3.2. Identification and sequence analysis of serine proteases

LocustDB, containing 12,161 unigenes (Ma et al., 2006), was searched using insect trypsins and chymotrypsins as query with 1e−15 as cutoff. In total, 27 single sequences and 68 preassembled contig sequences (for a total of 881 ESTs) with similarity to serine proteases were identified. After careful investigation, the sequences were assembled into 20 different transcripts which we are confident code for different serine proteases. For many EST sequences, very closely related transcripts were identified, showing only minor nucleotide changes. Since LocustDB was constructed using multiple locust larvae, it is likely that small nucleotide sequence differences reflect population variations (*i.e.* single nucleotide polymorphisms). For 16 transcripts, the coding sequence spans the entire mature protease, from the N-terminal activation site to a stop codon. The 4 other sequences lacked a predicted C-terminus. However, these sequences encode a substantial part of the mature protein. Therefore (with the exception of LmTry1A) the sequences were long enough to be included in the phylogenetic analysis. The predicted full length protease sequences range in size from 220 to 251 amino acids, corresponding to predicted molecular weights between 22.0 and 26.1 kDa. They have an average amino acid identity of 37.8%, ranging from 23.6 to 93.1% between members. An additional 13 predicted partial serine protease homologue sequences could be identified. However, their sequences are too small or lack important parts. Therefore, these will not be assessed here any further.

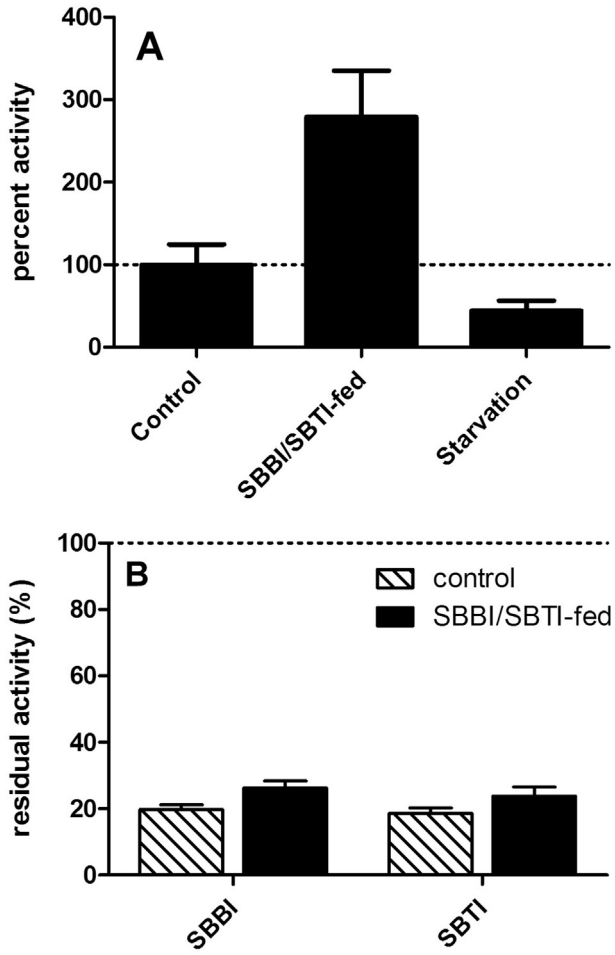

Fig. 2. Average proteolytic activity and sensitivity of secreted midgut enzymes from animals in different feeding conditions. (A) Percent activity of last instar larvae fed on SBBI/SBTI or starved for three days was normalized against midgut weight and compared to control animals ($n > 6$) ± SEM. (B) Inhibitor sensitivities to SBBI and SBTI of gut proteases derived from PI adapted and un-adapted animals were compared. For each inhibitor mean residual activity is represented ($n > 15$) ± SEM.

A multiple sequence alignment of the predicted amino acid sequences is presented in Fig. 3. All sequences share conserved amino acid motifs, including the N terminal activation site of the mature protein and the regions flanking the members of the catalytic triad, His57, Asp102 and Ser195 (bovine chymotrypsin numbering) (Table 1). Correct positioning in the active site of the tertiary structure is crucial for proteolytic activity (Hedstrom, 2002). The proteases were predicted to possess the core structural fold of the serine protease SA1 family, consisting of a C-terminal α-helix and two β-barrels, bridged by three conserved disulphide bridges: C42–C58, C168–C82 and C191–C220 (Várallyay et al., 1997). Interestingly, LmChy7 seems to have lost both cysteine residues from C42–C58 and LmChy8 is missing cysteine residue C220. The catalytic triad residues are well conserved in all transcripts with the exception of LmSPH1, where both His57 and Ser195 are replaced with Gln and Ile, respectively. Because of the lack of conservation in active site residues, it is very unlikely that LmSPH1 possesses chymotryptic activity; therefore, it was classified as serine protease homologue. With the exception of LmSPH1 and LmChy13, the GDSGG motif surrounding Ser195 is totally conserved in all sequences (Table 1). The characteristic TAAHC motif surrounding His57 is conserved in only 10 out of 20 proteases, while TAGHC is observed 6 times and SAAHC 3 times. The DIA motif is also less conserved, where variations include DIG, DVA and DFA.

Enzyme specificity is often attributed to the amino acid residue at position 189 (bovine chymotrypsin numbering). Trypsin-like enzymes possess a negatively charged Asp189 residue in their active site, while chymotrypsin-like enzymes utilize a small amino acid, such as Ser189 or Gly189 (Szabó et al., 2003). Based on this substrate specificity site, 6 of the identified serine proteases are designated as putative trypsin-like proteins, while the others encode putative chymotrypsin-like enzymes.

3.3. Phylogenetic analysis

An unrooted maximum likelihood tree (500 bootstrap repeats) of *L. migratoria* serine protease sequences is shown in Fig. 4. With the exception of LmTry1A all sequences were included. The phylogenetic analysis clearly shows the existence of three distinct groups of *L. migratoria* serine proteases that are all supported by high bootstrap values at the base. Cluster I contains all putative trypsin sequences (LmTry1B, 1C, 2A, 2B, 3) in addition to a number of putative chymotrypsin sequences (LmChy1, 2, 3, 4). Cluster II includes the putative chymotrypsin sequences LmChy5 and 6 and Cluster III contains the putative chymotrypsins LmChy7, 8, 9, 10, 11, 12, 13, as well as LmSPH1.

In a larger phylogenetic analysis, a maximum likelihood tree was constructed using additional trypsin and chymotrypsin sequences from over 60 different insect species belonging to a variety of insect orders (Diptera, Coleoptera, Hymenoptera, Siphonaptera, Hemiptera and Orthoptera) (Supplementary Data Fig. S3). With few exceptions, a clear division between putative insect trypsins and chymotrypsins is present, containing large monophyletic clusters of lepidopteran sequences and somewhat less clearly separated clusters from other insects. As expected, putative chymotrypsins of *L. migratoria* cluster III belong to the insect chymotrypsin group, while cluster I serine proteases are positioned within the subgroup of the trypsins. Interestingly, dipteran late trypsins cluster within the chymotrypsin subgroup. Surprisingly, LmChy5 and LmChy6 cluster together with this late trypsin gene subfamily. In addition, most trypsin-like sequences from Hymenoptera also cluster within the insect chymotrypsins.

3.4. Expression of serine proteases changes rapidly during different feeding conditions

To investigate the transcript levels of LmChy and LmTry in several tissues and different feeding conditions, (RT-)qPCR primers were designed for representative members of the three *L. migratoria* serine protease clusters. We developed primers that could discriminate between 3 of the identified trypsin-like sequences (LmTry1B, LmTry2A, LmTry2B), the chymotrypsin-like sequences from cluster I (LmChy1, LmChy2, LmChy3 and LmChy4), two members of cluster II (LmChy5 and LmChy6) and 3 members of cluster III (LmChy8, LmChy10 and LmChy13).

For this set of serine proteases, the transcript profile was determined in nine different tissues. A summary of the results is presented in Table 2. Expression of all genes was restricted to the midgut and caeca, with the exception of 4 genes (LmTry2A, LmTry2B, LmChy8, LmChy13) that had some minor additional expression in the foregut and very low expression levels of 2 genes (LmChy8, LmChy10) were observed in muscle tissue. For all genes, expression was highest in the caeca, expression levels being 5 to 150 times higher compared to midgut.

Transcript levels for these genes in both midgut and caeca were also quantified after a three-day starvation period (Fig. 5). Without any exception, expression levels declined strongly to around 20% of

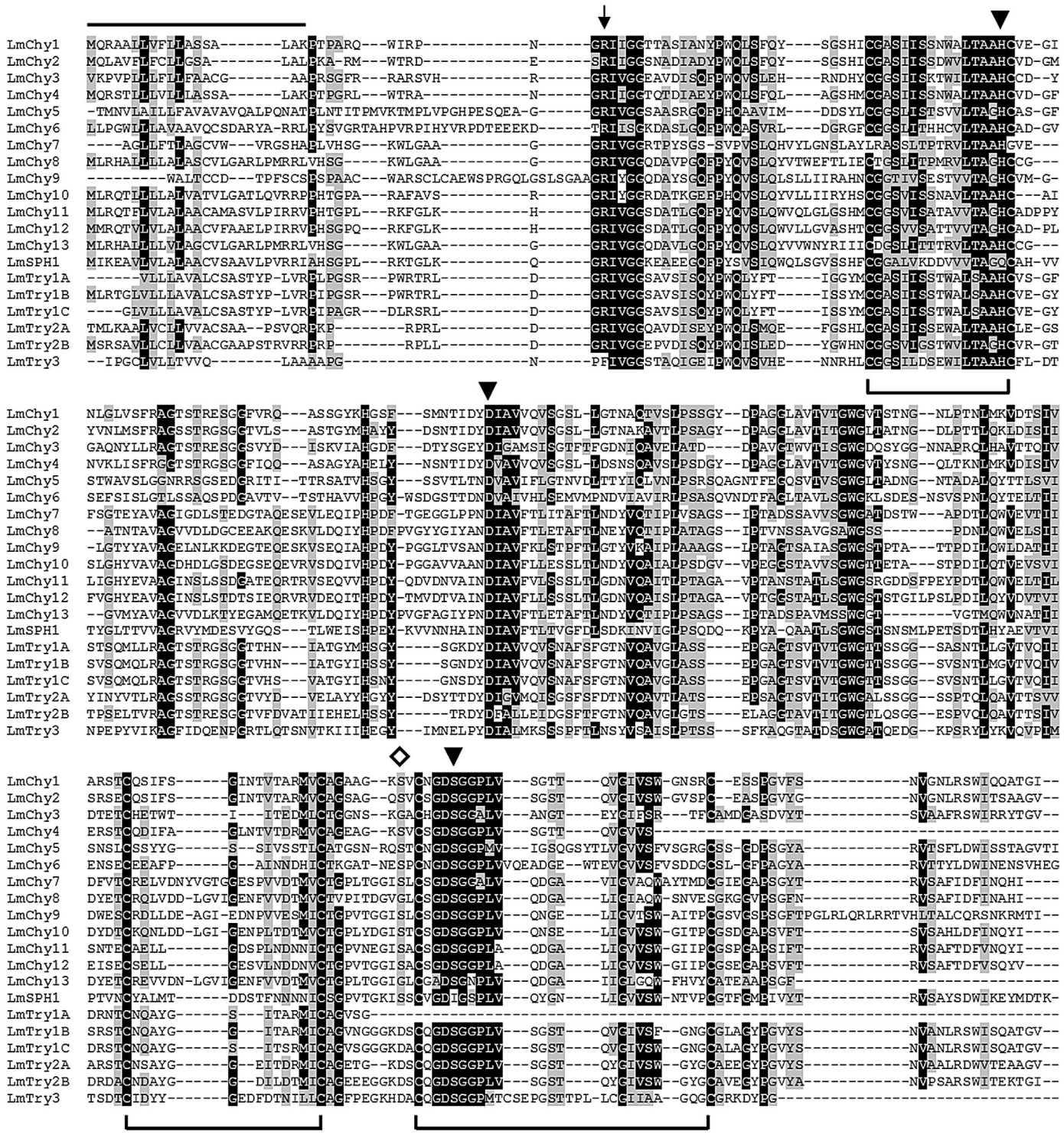

Fig. 3. Analysis of serine protease sequences from *L. migratoria*. A multiple sequence alignment of putative trypsin and chymotrypsin sequences was carried out using the MAFFT algorithm. Conservation threshold for shading is 65% identity. Identical residues are shaded black, similar residues are shaded gray. Predicted signal peptides are indicated by a line. The cleavage site generating the active protein is highlighted by an arrow. Indicated by a triangle at the top are residues that are part of the catalytic triad, His57, Asp102 and Ser195 (bovine trypsin numbering). The substrate determinant (residue 189, bovine trypsin numbering) is indicated by a diamond. Connecting disulfide bridges between cysteine residues are represented.

control levels. This strong reduction was observed in both parts of the digestive system.

To investigate possible PI induced expression differences, larvae were dissected at several time points after a single first meal containing inhibitors plus after prolonged exposure to an artificial diet containing SBBI and SBTI. Results show that expression levels from members of all three clusters changed rapidly. Four hours after ingesting SBBI and SBTI, protease expression was elevated in both midgut and gastric caeca (Fig. 6). Fold changes varied between 1.5 and 6 for the midgut and 1.5 and 3 for caeca, with LmTry2A,

Table 1
Summary of the characteristics and conserved motifs from identified cDNAs encoding putative serine proteases in *Locusta migratoria*.

Name	Genbank accession	Predicted mass (kDa)	Trypsinogen acitvation site	Conserved motifs around catalytic triad residues			AA189	Putative activity[a]	Cluster
				TAA*H*C	*D*IA	GDSGG			
LmTry1A[d]	BK008819	–	RIVGG	SAA*H*C	*D*IA	–	–	Trypsin[b]	I
LmTry1B	BK008820	22.0	RIVGG	SAA*H*C	*D*IA	GDSGG	D	Trypsin	I
LmTry1C	BK008821	22.0	RIVGG	SAA*H*C	*D*IA	GDSGG	D	Trypsin	I
LmTry2A	BK008822	22.9	RIVGG	TAA*H*C	*D*IG	GDSGG	D	Trypsin	I
LmTry2B	BK008823	23.2	RIVGG	TAG*H*C	*D*FA	GDSGG	D	Trypsin	I
LmTry3[d]	BK008824	–	FIVGG	TAA*H*C	*D*IA	GDSGG	D	Trypsin	I
LmChy1	BK008825	22.7	RIIGG	TAA*H*C	*D*IA	GDSGG	S	Chymotrypsin	I
LmChy2	BK008826	22.6	RIIGG	TAA*H*C	*D*IA	GDSGG	S	Chymotrypsin	I
LmChy3	BK008827	23.4	RIVGG	TAA*H*C	*D*IG	GDSGG	G	Chymotrypsin	I
LmChy4[d]	BK008828	–	RIIGG	TAA*H*C	*D*VA	GDSGG	S	Chymotrypsin	I
LmChy5	BK008829	23.9	RIVGG	TAG*H*C	*D*VA	GDSGG	S	Chymotrypsin	II
LmChy6	BK008830	25.0	RIISG	TAA*H*C	*D*VA	GDSGG	S	Chymotrypsin	II
LmChy7	BK008831	24.6	RIVGG	TAA*H*G	*D*IA	GDSGG	S	Chymotrypsin	III
LmChy8	BK008832	24.4	RIVGG	TAG*H*C	*D*IA	GDSGG	G	Chymotrypsin	III
LmChy9	BK008833	26.1	RIYGG	TAG*H*C	*D*IA	GDSGG	S	Chymotrypsin	III
LmChy10	BK008834	24.5	RIYGG	TAA*H*C	*D*IA	GDSGG	S	Chymotrypsin	III
LmChy11	BK008835	24.4	RIVGG	TAG*H*C	*D*IA	GDSGG	S	Chymotrypsin	III
LmChy12	BK008836	24.0	RIVGG	TAG*H*C	*D*IA	GDSGG	S	Chymotrypsin	III
LmChy13[d]	BK008837	–	RIVGG	TAA*H*C	*D*IA	ADSGN	G	Chymotrypsin	III
LmSPH1	BK008838	25.4	RIVGG	TAGQC	*D*IA	GDIGS	S	SPH[c]	III

Bold sequences indicate transcripts that were used in expression analysis.
[a] Putative activity based on substrate determinant AA189.
[b] Putative activity based on inference from sequence identity with LmTry1B and LmTry1A.
[c] SPH, serine protease homologue, missing active site residue (catalytic triad residues are italicized).
[d] Partial sequence with no stop codon, no predicted mass of mature enzyme possible.

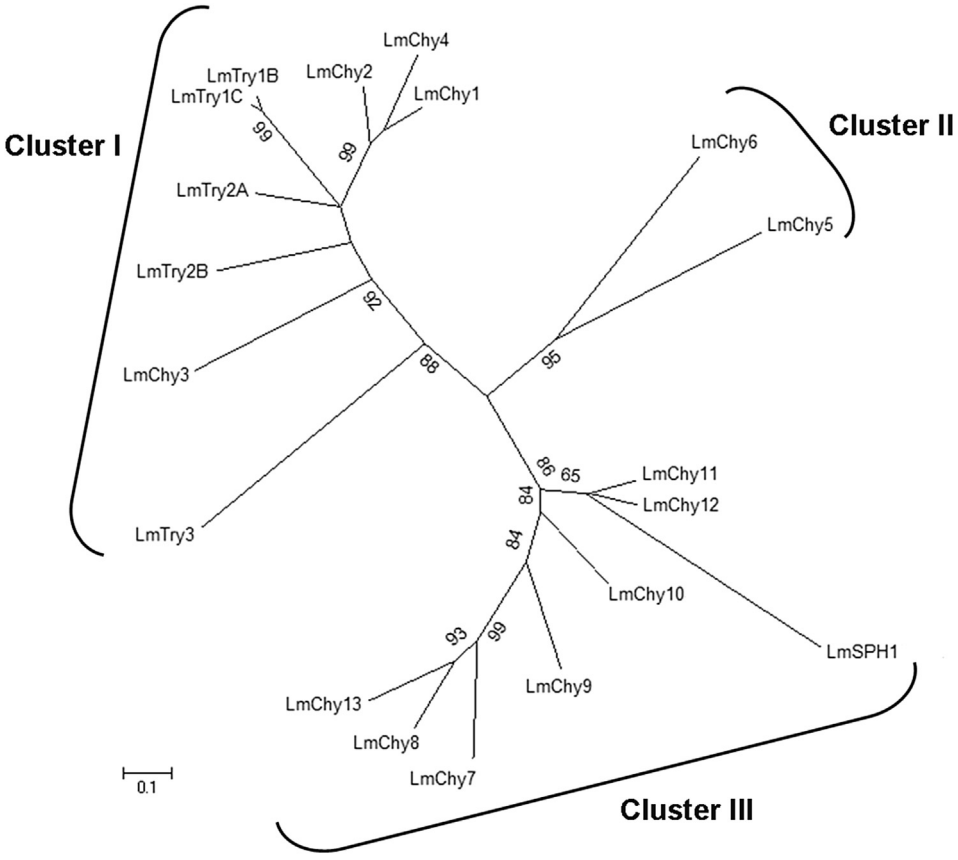

Fig. 4. Unrooted maximum likelihood phylogenetic tree of *L. migratoria* serine proteases. The concensus tree based on 500 bootstrap replicates is presented, showing three different clusters. Only bootstrap values higher than 60 are shown.

Table 2
Summary of relative abundance of protease transcript levels in different tissues of *Locusta migratoria*.

Name	MTb	CNS[b]	Fg	Mg[a]	Hg	Ca	Fb	Mu	Rs
LmChy1	–	–	–	1 ± 0.25	–	11 ± 2	–	–	–
LmChy2	–	–	–	1 ± 0.06	–	8 ± 1	–	–	–
LmChy3	–	–	–	1 ± 0.09	–	14 ± 3	–	–	–
LmChy4	–	–	–	1 ± 0.34	–	10 ± 2	–	–	–
LmTry1B	–	–	–	1 ± 0.46	–	18 ± 6	–	–	–
LmTry2A	–	–	0.11 ± 0. 3	1 ± 0.57	–	138 ± 20	–	–	–
LmTry2B	–	–	0.09 ± 0.02	1 ± 0.41	–	42 ± 3	–	–	–
LmChy5	–	–	–	1 ± 0.37	–	11 ± 1	–	–	–
LmChy6	–	–	–	1 ± 0.11	–	6 ± 2	–	–	–
LmChy8	–	–	0.20 ± 0.05	1 ± 0.17	–	152 ± 26	–	0.10 ± 0.03	–
LmChy10	–	–	–	1 ± 0.10	–	7 ± 1	–	0.05 ± 0.02	–
LmChy13	–	–	0.10 ± 0.04	1 ± 0.19	–	45 ± 9	–	–	–

MTb: Malpighian tubules, CNS: central nervous system, Fg: foregut, Mg: midgut, Hg: hindgut, Ca: caeca, Fb: fatbody, Mu: muscles, and Rs: reproductive system.
[a] Relative quantity for all tissues is compared to expression levels in midgut (RQ = 1). No detectable expression (RQ < 0.05) is indicated by –.
[b] CNS includes brain, optic lobes, corpora cardiaca and corpora allata.

LmTry2B and LmChy8 being the most differentially expressed. Interestingly, only members of cluster I maintained higher transcript levels after exposure to PI for two consecutive days. In addition, expression levels of all proteases returned to control levels approximately 12 h after larvae were no longer allowed to keep feeding on diet containing SBBI and SBTI (data not shown).

4. Discussion

We confirmed the use of the serine proteases trypsin and chymotrypsin as major protein-digesting enzymes in *L. migratoria*, in addition to a minor contribution of cysteine proteases to the total proteolytic activity in the gut. Although SBBI and SBTI proved to be effective inhibitors of the serine protease activity in gut secretions of *L. migratoria*, larvae were able to quickly adjust their midgut physiology after ingestion of these inhibitors. These results are in agreement with our previous findings that locusts are capable to survive on a diet containing high doses of protease inhibitors, as was shown for *S. gregaria* (Spit et al., 2012). No effects on larval development were observed, while proteolytic activity was increased more than twofold after three days, similar to what was observed in *S. gregaria* when equal inhibitor concentrations were used (Spit et al., 2012). Starvation of locusts did not result in elevated levels of proteolytic activity. On the contrary, a strong decrease in activity was observed. The latter presumably represents an energy saving mechanism and prevents self-digestion of the midgut by an excess amount of proteases in the digestive system when no food is present. Drastically reduced protease activity is commonly observed in other insects, for example in starved caterpillars (Bown et al., 1997; Broehan et al., 2008; Zhan et al., 2011). Our data indicate that the rise in protease activity followed upon PI ingestion is not simply due to nutrient deprivation, but operates via another mechanism, that has yet to be identified. Indeed, other authors have also reported that supplying free amino acids in combination with protease inhibitors did not prevent PI induced response mechanisms, indicating that the presence of protease inhibitors in the diet was not signaled through a lack of free amino acids (Bown et al., 2004).

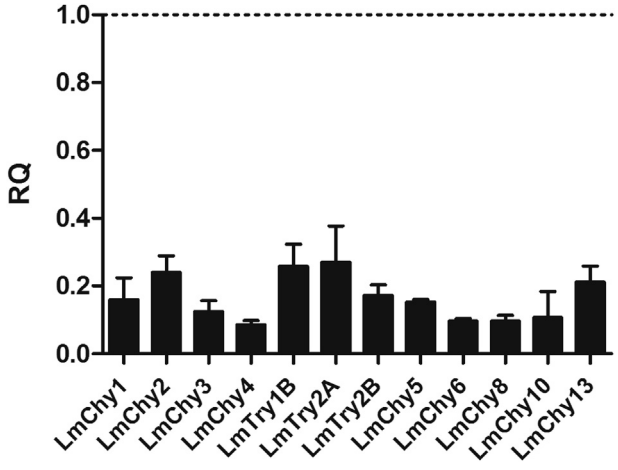

Fig. 5. Starvation induces downregulation of *L. migratoria* serine protease expression. Relative quantity of transcript is shown after three days of starvation. Relative transcript levels are calculated for each protease separately compared to a control condition (RQ = 1). Expression levels between different genes cannot be compared. Data were normalized against two reference genes, RP49 and RPS13. Means ± SEM are presented for each transcript tested ($n \geq 3$ pools, 5 individuals per pool).

This study is the first to describe the identification of serine protease transcript sequences in *L. migratoria*. At least 20 different serine protease transcript sequences could be identified, which clustered in three phylogenetically distinct groups. There was substantial redundancy in transcripts encoding these proteases. In addition, several ESTs encoding only partial proteases were also found, suggesting that other proteases might still be present. Based on our analysis, we are confident that the serine proteases that contribute most to digestion are represented in this survey, since these will have the highest numbers of ESTs encoding them, thus allowing for the assembly of their coding sequence. The existence of large numbers of serine protease cDNAs is not surprising, since with the progress in genomic and transcriptomic technology, a growing number of insects are found to express large repertoires of proteolytic enzymes (Ge et al., 2012; George et al., 2008; Marshall et al., 2008; Mazumdar-Leighton et al., 2000; Mazumdar-Leighton and Broadway, 2001a,b; Prabhakar et al., 2007; Simpson et al., 2007; Vinokurov et al., 2006; Zhu and Baker, 1999). Expression for all tested *L. migratoria* transcripts was almost exclusively restricted to the digestive system, indicating a role in the digestion of dietary protein. The presence of a secretion signal peptide and the fact that these transcripts are downregulated during starvation further support this idea.

Small variations in conserved motifs surrounding catalytic triad residues are observed, in agreement with previous reports made in other insects (Coates et al., 2006; Marshall et al., 2008; Prabhakar et al., 2007; Simpson et al., 2007; Zhu and Baker, 1999). However, since the effects of these sequence variations on enzyme activity and sensitivity remain unclear, these results emphasize the need for further biochemical studies to characterize the implications of aberrations in the amino acid sequences of various proteases. At

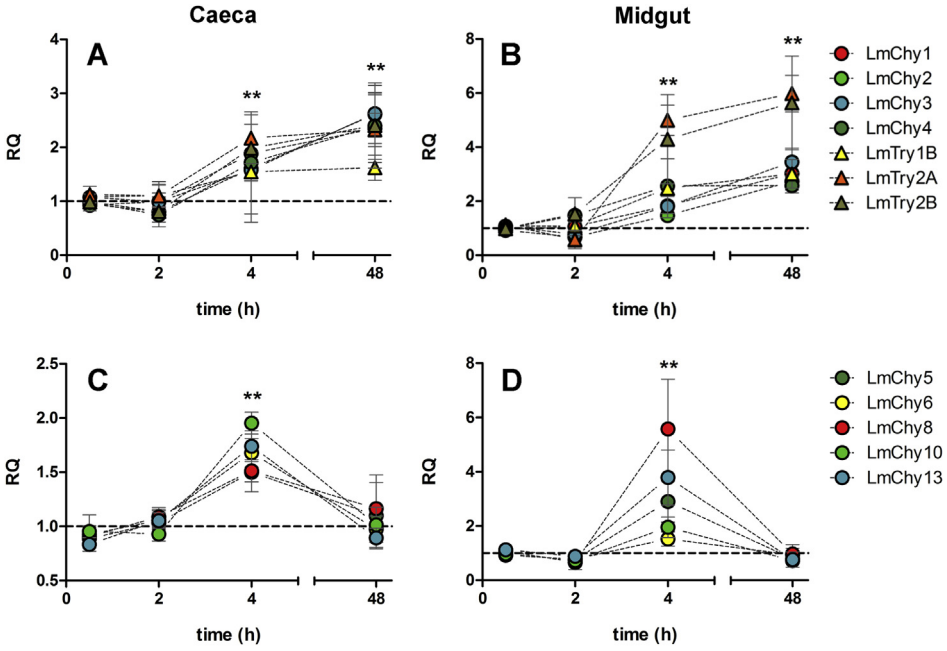

Fig. 6. Expression of proteases in the digestive system of *L. migratoria* after feeding on a diet containing SBBI and SBTI. Relative expression of serine protease members from cluster I (A and B) and clusters II and III (C,D) in caeca (A and C) and midgut (B and D) at different time points after transfer to an artificial diet containing SBBI/SBTI. Means ± SEM are calculated against control values (RQ = 1) for each gene and time point separately ($n \geq 3$, 5 individuals per pool). Levels of significance (ANOVA) are indicated by ** for $p < 0.01$.

least one of the expressed full length transcripts is predicted to be inactive based on alterations in the catalytic triad residues (LmSPH1). Furthermore, two transcripts (LmChy7 and LmChy 8) were found to be missing one or more cysteine residues that are involved in disulphide bridge formation. The implications for enzyme activity are unclear, but, in general, three disulphide bridges are expected to be important for maintaining the tertiary structure of the enzyme and thus proteolytic activity. Expression of putative inactive enzymes in the digestive system of insects is a reoccurring phenomenon and was also observed in for example *Costelytra zealandica* (Marshall et al., 2008), *Chilo suppressalis* (Ge et al., 2012), *Helicoverpa armigera* (Bown et al., 1997; Mazumdar-Leighton et al., 2000), *S. gregaria* (Chiou et al., 1998) and *Epiphyas postvittana* (Simpson et al., 2007). It has been suggested that production of inactive proteases could sequester dietary ingested protease inhibitors, and so provide an adaptive advantage without the need for excessive amounts of protease activity in the midgut (Christeller, 2005; Mazumdar-Leighton et al., 2000).

Our phylogenetic analysis indicates a division between insect trypsin and chymotrypsin-like sequences, which is in line with observations from several other reports (Hughes and Vogler, 2006; Lopes et al., 2004; Marshall et al., 2008; Oliveira-Neto et al., 2004). Sequences from *L. migratoria* were found to be divided into three distinct clusters. Interestingly, cluster I was positioned within the insect trypsin subgroup but contained both trypsin- and chymotrypsin-like sequences, suggesting that the latter might have been derived from an ancestral trypsin-like sequence and subsequently evolved chymotrypsin-like substrate specificity. Another noticeable fact was the position of a hymenopteran trypsin cluster and a late trypsin cluster within the insect chymotrypsins. The latter had already been observed once in a previously reported phylogenetic analysis of insect serine proteases, suggesting that this subfamily might have been derived from a chymotrypsin ancestor (Marshall et al., 2008). Careful investigation of the late trypsin amino acid sequences at the substrate determinant position demonstrated that 6 of these possess a Ser189, while 2 other sequences code for an Asn189. None of them contains the characteristic Asp189 that specifies other trypsin-like enzymes. As most members were only identified *in silico*, it would be very interesting to find out if these late trypsins indeed are active as trypsin-like enzymes or instead possess chymotrypsin activity. Surprisingly, LmChy5 and LmChy6 from *L. migratoria* appeared to be part of this late trypsin gene subfamily in our analysis. If the observed clustering is not the result of a phylogenetic anomaly, close clustering of these sequences is remarkable, since there is a relatively large evolutionary distance between Orthoptera and Diptera and late trypsin-like sequences had not been described in any other insect orders except for Diptera until now.

The transcript profiling data do not allow us to predict the absolute quantities of the corresponding proteolytic enzymes in the digestive system, nor can we directly compare expression levels between different proteases. However, it is likely that only a small proportion of these genes is highly expressed at the active protein level in the digestive system. For 7 identified transcripts, the translated N-terminal fragment is identical to the N-terminus of five known serine proteases that were previously purified from the midgut of *L. migratoria* (LmChy1, LmChy2, LmTry1, LmTry2A and LmTry2B) (Lam et al., 1999, 2000). For LmTry1, three different isoforms could be identified with exactly the same N-terminal sequence. These sequences were designated LmTry1A, LmTry1B and LmTry1C accordingly. All the transcript sequences corresponding to purified proteases belong to cluster I, indicating that members of this cluster may constitute the major digestive proteases of *L. migratoria*. Nevertheless, the observed large diversity in transcripts likely permits a great flexibility in substrate specificity when feeding on different host plants or when encountering different dietary protease inhibitors.

Upregulation of serine protease expression after PI ingestion has been well documented for some insects, mostly generalist feeders. The most common reports include the overproduction of identical, PI sensitive proteases (Brioschi et al., 2007; De Leo et al., 1998; Markwick et al., 1998) or the expression of new proteases that are insensitive to the protease inhibitors that were applied (Bown et al., 1997, 2004; Brito et al., 2001; De Oliveira et al., 2013). Although it

cannot be excluded that certain PI insensitive enzymes exist in *L. migratoria*, our current work shows that the observed physiological adaptation to SBBI and SBTI was primarily based on the overexpression of existing PI sensitive enzymes, rather than on the '*de novo*' production of other PI insensitive proteases, as evidenced by an unchanged PI sensitivity of the induced proteolytic activity. Upregulation of all tested serine protease transcripts occurred rapidly, 4 h after PI ingestion. Results indicate that the PI induced compensation response in *L. migratoria* followed a distinct pattern of initial upregulation of all tested proteases, subsequently followed by a phase, wherein only expression levels of members of cluster I remained elevated. These observations correlate perfectly with the observed increase in proteolytic activity after 48 h of feeding on a diet containing SBBI and SBTI. A similar temporal pattern was observed in *H. armigera*, where chronic ingestion of SBTI also resulted in an initial upregulation of all tested proteases, but was succeeded by the downregulation of the enzymes most susceptible to SBTI, ultimately resulting in an activity change toward SBTI insensitive proteases (Bown et al., 2004). However, in our study, no apparent activity change toward insensitive enzymes was observed. We hypothesize that the observed upregulation of represents a general mechanism to rapidly increase the wide repertoire of proteolytic enzyme activities in the gut in response to PI in the diet, while cluster I genes, which may encode the enzymes that contribute most to protein digestion, remain upregulated after prolonged exposure times, allowing for the maintenance of a sufficiently high protein-degrading enzyme activity in the midgut. However, the exact molecular mechanisms by which insects sense and react to protease inhibitors in their diet remain elusive.

Insect adaptation to PI has been studied most extensively in members of the Lepidoptera and Coleoptera (Ahn et al., 2007; Bayés et al., 2006; Brioschi et al., 2007; Govind et al., 2010; Jongsma et al., 1995; Mazumdar-Leighton and Broadway, 2001b; Moon et al., 2004; Petek et al., 2012; Volpicella et al., 2002; Zhu-Salzman et al., 2003). It is interesting to notice that Orthoptera are now also shown to possess similar response mechanisms for the physiological adaptation to the presence of (plant) protease inhibitors in their diet, suggesting that these mechanisms, which are expected to be crucial for animal fitness and survival, were already present early in insect evolution. It is tempting to assume that general upregulation of proteolytic enzymes is a more primitive response to PI ingestion, whereas the expression of an alternative set of insensitive enzymes is a more sophisticate response that is observed only in more derived insects.

Acknowledgments

The authors explicitly thank Roger Jonckers for maintaining the locust culture and Joost Van Duppen for technical assistance. J.S., M.H. and S.Z. were supported by a PhD fellowship of the IWT (Agency for Innovation by Science and Technology). In addition, the authors also gratefully acknowledge the KU Leuven Research Foundation (GOA/11/02) and the Research Foundation of Flanders (FWO) (FWO-G031112N) for financial support.

Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.ibmb.2014.03.002.

References

Ahn, J.-E., Lovingshimer, M.R., Salzman, R.A., Presnail, J.K., Lu, A.L., Koiwa, H., Zhu-Salzman, K., 2007. Cowpea bruchid *Callosobruchus maculatus* counteracts dietary protease inhibitors by modulating propeptides of major digestive enzymes. Insect Mol. Biol. 16, 295–304.

Badisco, L., Huybrechts, J., Simonet, G., Verlinden, H., Marchal, E., Huybrechts, R., Schoofs, L., De Loof, A., Vanden Broeck, J., 2011. Transcriptome analysis of the desert locust central nervous system: production and annotation of a *Schistocerca gregaria* EST database. PLoS One 6, e17274.

Bayés, A., de la Vega, M.R., Vendrell, J., Aviles, F.X., Jongsma, M.A., Beekwilder, J., 2006. Response of the digestive system of *Helicoverpa zea* to ingestion of potato carboxypeptidase inhibitor and characterization of an uninhibited carboxypeptidase B. Insect Biochem. Mol. Biol. 36, 654–664.

Bown, D.P., Wilkinson, H.S., Gatehouse, J.A., 1997. Differentially regulated inhibitor-sensitive and insensitive protease genes from the phytophagous insect pest, *Helicoverpa armigera*, are members of complex multigene families. Insect Biochem. Mol. Biol. 27, 625–638.

Bown, D.P., Wilkinson, H.S., Gatehouse, J.A., 2004. Regulation of expression of genes encoding digestive proteases in the gut of a polyphagous lepidopteran larva in response to dietary protease inhibitors. Physiol. Entomol. 29, 278–290.

Brioschi, D., Nadalini, L.D., Bengtson, M.H., Sogayar, M.C., Moura, D.S., Silva-Filho, M.C., 2007. General up regulation of *Spodoptera frugiperda* trypsins and chymotrypsins allows its adaptation to soybean proteinase inhibitor. Insect Biochem. Mol. Biol. 37, 1283–1290.

Brito, L.O., Lopes, A.R., Parra, J.R., Terra, W.R., Silva-Filho, M.C., 2001. Adaptation of tobacco budworm *Heliothis virescens* to proteinase inhibitors may be mediated by the synthesis of new proteinases. Comp. Biochem. Physiol. B. Biochem. Mol. Biol. 128, 365–375.

Broehan, G., Kemper, M., Driemeier, D., Vogelpohl, I., Merzendorfer, H., 2008. Cloning and expression analysis of midgut chymotrypsin-like proteinases in the tobacco hornworm. J. Insect Physiol. 54, 1243–1252.

Chiou, S.J., Vanden Broeck, J., Janssen, I., Borovsky, D., Vandenbussche, F., Simonet, G., De Loof, a, 1998. Cloning of the cDNA encoding Scg-SPRP, an unusual Ser-protease-related protein from vitellogenic female desert locusts (*Schistocerca gregaria*). Insect Biochem. Mol. Biol. 28, 801–808.

Christeller, J.T., 2005. Evolutionary mechanisms acting on proteinase inhibitor variability. FEBS J. 272, 5710–5722.

Coates, B.S., Hellmich, R.L., Lewis, L.C., 2006. Sequence variation in trypsin- and chymotrypsin-like cDNAs from the midgut of *Ostrinia nubilalis*: methods for allelic differentiation of candidate *Bacillus thuringiensis* resistance genes. Insect Mol. Biol. 15, 13–24.

De Leo, F., Bonade-Bottino, M., Ceci, L., Gallerani, R., Jouanin, L., 1998. Opposite effects on *Spodoptera littoralis* larvae of high expression level of a trypsin proteinase inhibitor in transgenic plants. Plant Physiol. 118, 997–1004.

De Oliveira, C.F.R., de Paula Souza, T., Parra, J.R.P., Marangoni, S., Silva-Filho, M.D.C., Macedo, M.L.R., 2013. Insensitive trypsins are differentially transcribed during *Spodoptera frugiperda* adaptation against plant protease inhibitors. Comp. Biochem. Physiol. B. Biochem. Mol. Biol. 165, 19–25.

Gatehouse, J.A., 2011. Prospects for using proteinase inhibitors to protect transgenic plants against attack by herbivorous insects. Curr. Protein Pept. Sci. 12, 409–416.

Gatehouse, L.N., Shannon, A.L., Burgess, E.P., Christeller, J.T., 1997. Characterization of major midgut proteinase cDNAs from *Helicoverpa armigera* larvae and changes in gene expression in response to four proteinase inhibitors in the diet. Insect Biochem. Mol. Biol. 27, 929–944.

Ge, Z., Wan, P., Han, Z., 2012. Cloning and characterization of trypsin-and chymotrypsin-like genes in the striped rice stem borer, *Chilo suppressalis*. Genome 55, 281–288.

George, D., Ferry, N., Back, E., Gatehouse, A.M.R., 2008. Characterisation of midgut digestive proteases from the maize stem borer *Busseola fusca*. Pest Manag. 64 (1158), 1151–1158.

Govind, G., Mittapalli, O., Griebel, T., Allmann, S., Böcker, S., Baldwin, I.T., 2010. Unbiased transcriptional comparisons of generalist and specialist herbivores feeding on progressively defenseless *Nicotiana attenuata* plants. PLoS One 5, e8735.

Hedstrom, L., 2002. Serine protease mechanism and specificity. Chem. Rev. 102, 4501–4524.

Hughes, J., Vogler, A.P., 2006. Gene expression in the gut of keratin-feeding clothes moths (Tineola) and keratin beetles (Trox) revealed by subtracted cDNA libraries. Insect Biochem. Mol. Biol. 36, 584–592.

Jongsma, M.A., Bakker, P.L., Peters, J., Bosch, D., Stiekema, W.J., 1995. Adaptation of *Spodoptera exigua* larvae to plant proteinase inhibitors by induction of gut proteinase activity insensitive to inhibition. Proc. Natl. Acad. Sci. U. S. A 92, 8041–8045.

Kang, L., Chen, X., Zhou, Y., Liu, B., Zheng, W., Li, R., Wang, J., Yu, J., 2004. The analysis of large-scale gene expression correlated to the phase changes of the migratory locust. Proc. Natl. Acad. Sci. U. S. A 101, 17611–17615.

Katoh, K., Kuma, K., Toh, H., Miyata, T., 2005. MAFFT version 5: improvement in accuracy of multiple sequence alignment. Nucleic Acids Res. 33, 511–518.

Lam, W., Coast, G.M., Rayne, R.C., 1999. Isolation and characterisation of two chymotrypsins from the midgut of *Locusta migratoria*. Insect Biochem. Mol. Biol. 29, 653–660.

Lam, W., Coast, G.M., Rayne, R.C., 2000. Characterisation of multiple trypsins from the midgut of *Locusta migratoria*. Insect Biochem. Mol. Biol. 30, 85–94.

Livak, K.J., Schmittgen, T.D., 2001. Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta CT}$ method. Methods 25, 402–408.

Lopes, A.R., Juliano, M.A., Juliano, L., Terra, W.R., 2004. Coevolution of insect trypsins and inhibitors. Arch. Insect Biochem. Physiol. 55, 140–152.

Ma, Z., Yu, J., Kang, L., 2006. LocustDB: a relational database for the transcriptome and biology of the migratory locust (*Locusta migratoria*). BMC Genom. 7, 11.

Markwick, N.P., Laing, W.A., Christeller, J.T., McHenry, J.Z., Newton, M.R., 1998. Overproduction of digestive enzymes compensates for inhibitory effects of protease and alpha-amylase inhibitors fed to three species of leafrollers (Lepidoptera: Tortricidae). J. Econ. Entomol. 91, 1265–1276.

Marshall, S.D.G., Gatehouse, L.N., Becher, S.A., Christeller, J.T., Gatehouse, H.S., Hurst, M.R.H., Boucias, D.G., Jackson, T.A., 2008. Serine proteases identified from a *Costelytra zealandica* (White) (Coleoptera: Scarabaeidae) midgut EST library and their expression through insect development. Insect Mol. Biol. 17, 247–259.

Mazumdar-Leighton, S., Babu, C.R., Bennett, J., 2000. Identification of novel serine proteinase gene transcripts in the midguts of two tropical insect pests, *Scirpophaga incertulas* (Wk.) and *Helicoverpa armigera* (Hb.). Insect Biochem. Mol. Biol. 30, 57–68.

Mazumdar-Leighton, S., Broadway, R.M., 2001a. Transcriptional induction of diverse midgut trypsins in larval *Agrotis ipsilon* and *Helicoverpa zea* feeding on the soybean trypsin inhibitor. Insect Biochem. Mol. Biol. 31, 645–657.

Mazumdar-Leighton, S., Broadway, R.M., 2001b. Identification of six chymotrypsin cDNAs from larval midguts of *Helicoverpa zea* and *Agrotis ipsilon* feeding on the soybean (Kunitz) trypsin inhibitor. Insect Biochem. Mol. Biol. 31, 633–644.

Mello, M., Silva-Filho, M., 2002. Plant-insect interactions: an evolutionary arms race between two distinct defense mechanisms. Braz. J. Plant Physiol. 14, 71–81.

Moon, J., Salzman, R.A., Ahn, J.-E., Koiwa, H., Zhu-Salzman, K., 2004. Transcriptional regulation in cowpea bruchid guts during adaptation to a plant defence protease inhibitor. Insect Mol. Biol. 13, 283–291.

Oliveira-Neto, O.B., Batista, J.A.N., Rigden, D.J., Fragoso, R.R., Silva, R.O., Gomes, E.A., Franco, O.L., Dias, S.C., Cordeiro, C.M.T., Monnerat, R.G., Grossi-De-Sá, M.F., 2004. A diverse family of serine proteinase genes expressed in cotton boll weevil (*Anthonomus grandis*): implications for the design of pest-resistant transgenic cotton plants. Insect Biochem. Mol. Biol. 34, 903–918.

Pedra, J.H.F., Brandt, A., Westerman, R., Lobo, N., Li, H.-M., Romero-Severson, J., Murdock, L.L., Pittendrigh, B.R., 2003. Transcriptome analysis of the cowpea weevil bruchid: identification of putative proteinases and alpha-amylases associated with food breakdown. Insect Mol. Biol. 12, 405–412.

Petek, M., Turnšek, N., Gašparič, M.B., Novak, M.P., Gruden, K., Slapar, N., Popovič, T., Strukelj, B., Jongsma, M.A., 2012. a complex of genes involved in adaptation of *Leptinotarsa decemlineata* larvae to induced potato defense. Arch. Insect Biochem. Physiol. 79, 153–181.

Petersen, T.N., Brunak, S., von Heijne, G., Nielsen, H., 2011. SignalP 4.0: discriminating signal peptides from transmembrane regions. Nat. Methods 8, 785–786.

Prabhakar, S., Chen, M.-S., Elpidina, E.N., Vinokurov, K.S., Smith, C.M., Marshall, J., Oppert, B., 2007. Sequence analysis and molecular characterization of larval midgut cDNA transcripts encoding peptidases from the yellow mealworm, *Tenebrio molitor* L. Insect Mol. Biol. 16, 455–468.

Simpson, R.M., Newcomb, R.D., Gatehouse, H.S., Crowhurst, R.N., Chagné, D., Gatehouse, L.N., Markwick, N.P., Beuning, L.L., Murray, C., Marshall, S.D., Yauk, Y.-K., Nain, B., Wang, Y.-Y., Gleave, A.P., Christeller, J.T., 2007. Expressed sequence tags from the midgut of *Epiphyas postvittana* (Walker) (Lepidoptera: Tortricidae). Insect Mol. Biol. 16, 675–690.

Spit, J., Breugelmans, B., van Hoef, V., Simonet, G., Zels, S., Vanden Broeck, J., 2012. Growth-inhibition effects of pacifastin-like peptides on a pest insect: the desert locust, *Schistocerca gregaria*. Peptides 34, 251–257.

Szabó, E., Venekei, I., Böcskei, Z., Náray-Szabó, G., Gráf, L., 2003. Three dimensional structures of S189D chymotrypsin and D189S trypsin mutants: the effect of polarity at site 189 on a protease-specific stabilization of the substrate-binding site. J. Mol. Biol. 331, 1121–1130.

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., Kumar, S., 2011. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol. Biol. Evol. 28, 2731–2739.

Vandesompele, J., De Preter, K., Pattyn, F., Poppe, B., Van Roy, N., De Paepe, A., Speleman, F., 2002. Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. Genome Biol. 3. RESEARCH0034.1–RESEARCH0034.11.

Várallyay, E., Lengyel, Z., Gráf, L., Szilágyi, L., 1997. The role of disulfide bond C191-C220 in trypsin and chymotrypsin. Biochem. Biophys. Res. Commun. 230, 592–596.

Verlinden, H., Badisco, L., Marchal, E., Van Wielendaele, P., Vanden Broeck, J., 2009. Endocrinology of reproduction and phase transition in locusts. Gen. Comp. Endocrinol. 162, 79–92.

Vinokurov, K.S., Elpidina, E.N., Oppert, B., Prabhakar, S., Zhuzhikov, D.P., Dunaevsky, Y.E., Belozersky, M.A., 2006. Diversity of digestive proteinases in *Tenebrio molitor* (Coleoptera: Tenebrionidae) larvae. Comp. Biochem. Physiol. Part B Biochem. Mol. Biol. 145, 126–137.

Volpicella, M., Ceci, L.R., Cordewener, J., America, T., Gallerani, R., Bode, W., Jongsma, M.A., Beekwilder, J., 2002. Properties of purified gut trypsin from *Helicoverpa zea*, adapted to proteinase inhibitors. Eur. J. Biochem 270, 10–19.

Wei, Z., Yin, Y., Zhang, B., Wang, Z., Peng, G., Cao, Y., Xia, Y., 2007. Cloning of a novel protease required for the molting of *Locusta migratoria* manilensis. Dev. Growth Differ. 49, 611–621.

Zhan, Q., Zheng, S., Feng, Q., Liu, L., 2011. A midgut-specific chymotrypsin cDNA (Slctlp1) from *Spodoptera litura*: cloning, characterization, localization and expression analysis. Arch. Insect Biochem. Physiol. 76, 130–143.

Zhang, Z., Peng, Z.-Y., Yi, K., Cheng, Y., Xia, Y., 2012. Identification of representative genes of the central nervous system of the locust, *Locusta migratoria* manilensis by deep sequencing. J. Insect Sci. 12, 86.

Zhu, Y.C., Baker, J.E., 1999. Characterization of midgut trypsin-like enzymes and three trypsinogen cDNAs from the lesser grain borer, *Rhyzopertha dominica* (Coleoptera: Bostrichidae). Insect Biochem. Mol. Biol. 29, 1053–1063.

Zhu-Salzman, K., Koiwa, H., Salzman, R.A., Shade, R.E., Ahn, J.-E., 2003. Cowpea bruchid *Callosobruchus maculatus* uses a three-component strategy to overcome a plant defensive cysteine protease inhibitor. Insect Mol. Biol. 12, 135–145.